

Maskininlärning för väderprognoser (MLWP)

Ensembleprognoser och regional modellering

Martin Andrae

Avdelningen för Statistik och maskininlärning

Linköpings universitet

Vem är jag?

- BSc i Teknisk Fysik och MSc i Tillämpad Matematik från KTH
- Masteruppsats på SMHI:
 - *Väderprognosensembler med hjälp av generativa modeller*
 - Handledd av Tomas Landelius
- Doktorand i Statistik på LiU
 - 3 doktorander som forskar på ML/AI för väder

Innehåll

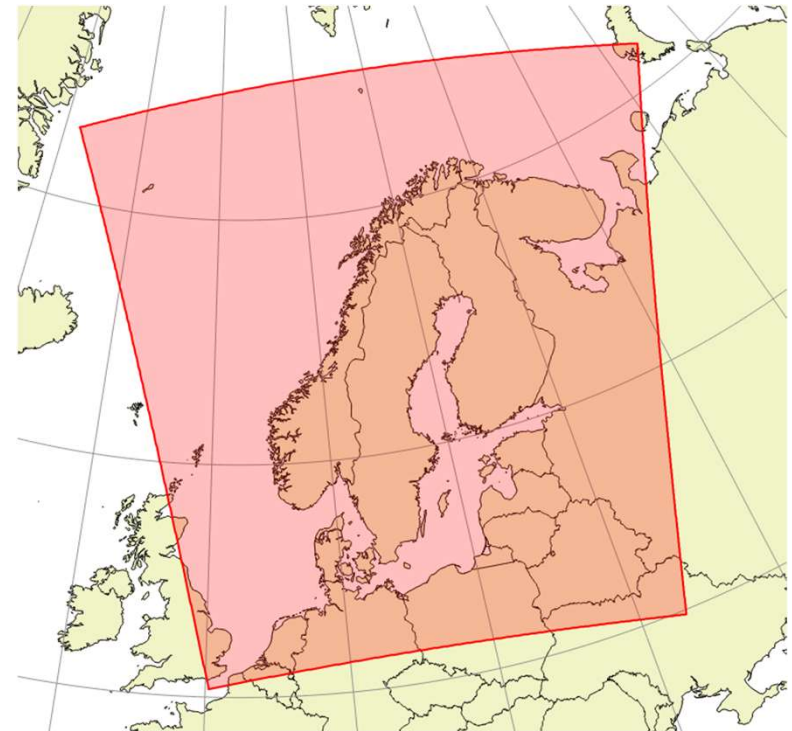
- Regional modellering
- Ensembleprognoser

Regional modellering

- Hur gör vi högupplösta prognoser över lokala områden?
- Hur hanterar vi randvillkor och global information?

Exempel: Neural-LAM ¹

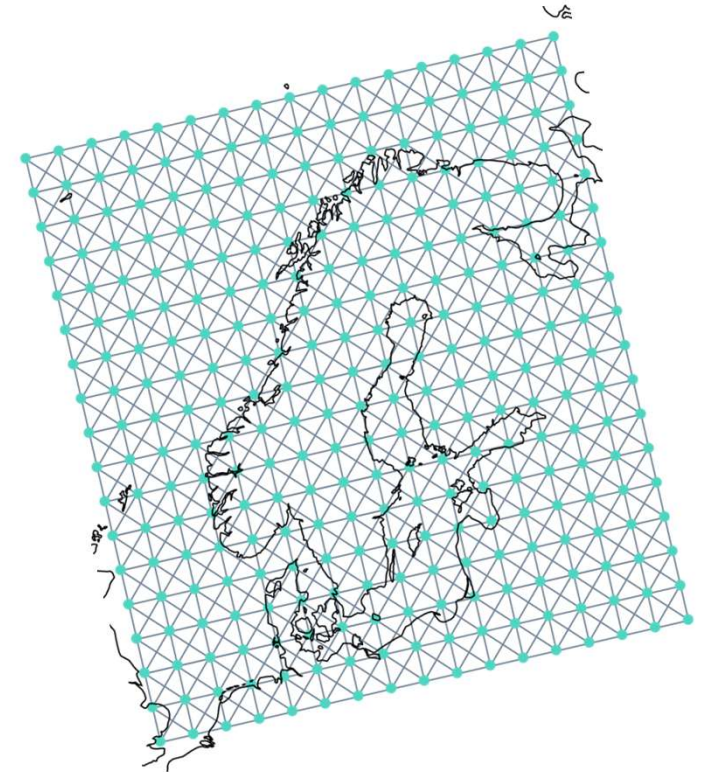
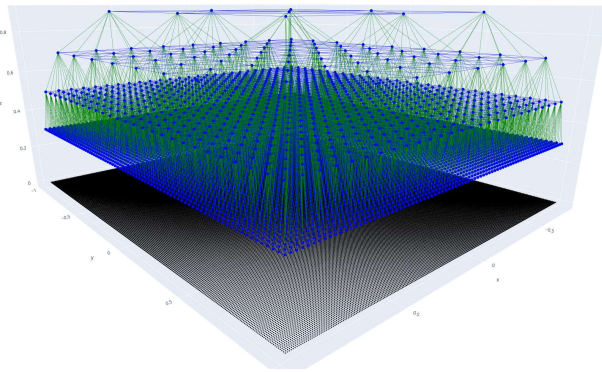
- Data från MEPS
- 10 km upplösning
- 57h prognoser med 3h steg
- Urval av variabler: Tryck, geopotential, vind, temperatur etc
- **Mål:** Kan vi emulera MEPS med ML?



¹ J. Oskarsson, et al. (2023). Graph-based Neural Weather Prediction for Limited Area Modeling. *Tackling Climate Change with Machine Learning: workshop at NeurIPS 2023*.

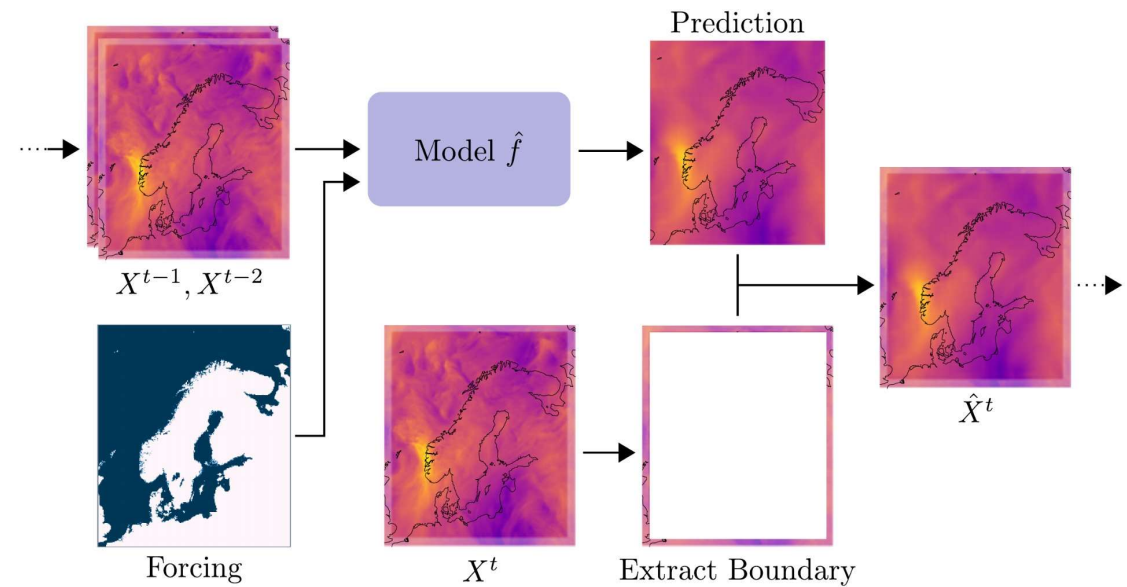
Grafiska Neurala Nätverk

- Används i Globala MLWP modeller
- Utnyttjar strukturen hos datan
- Lär sig lokala och globala samband



Randvillkor

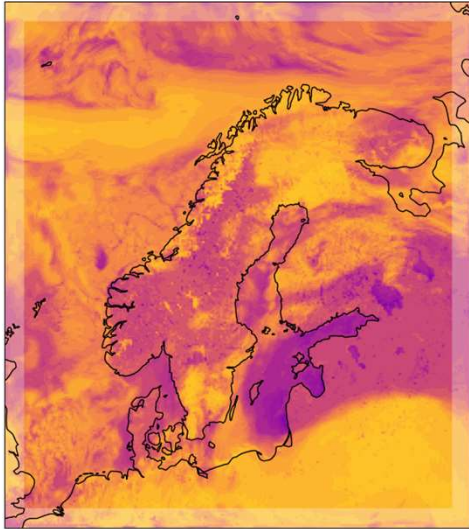
- **Här:** Exakta randvillkor från MEPS
- **Alternativ:** Exakta randvillkor från global modell



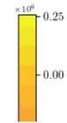
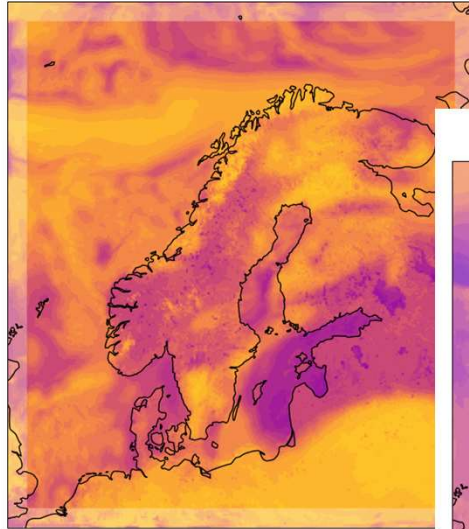
Exempelprognoser

nlwrs_0 (W/m²), t=1 (3 h)

Ground Truth



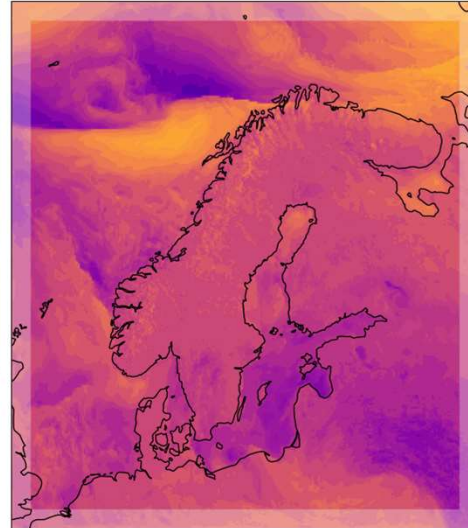
Prediction



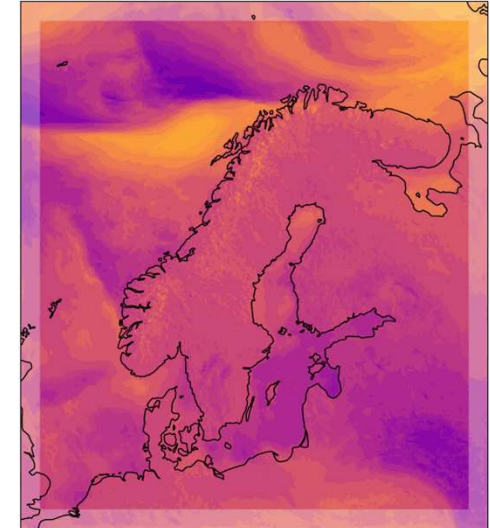
Terrestrisk strålning

u_65 (m/s), t=1 (3 h)

Ground Truth



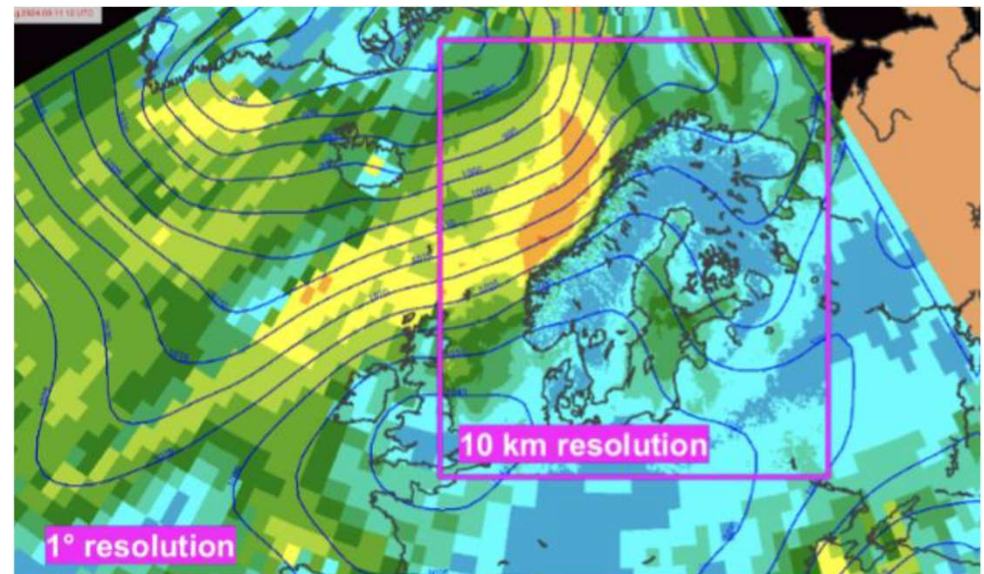
Prediction



U-komponent vind

Exempel: MET Norge ²

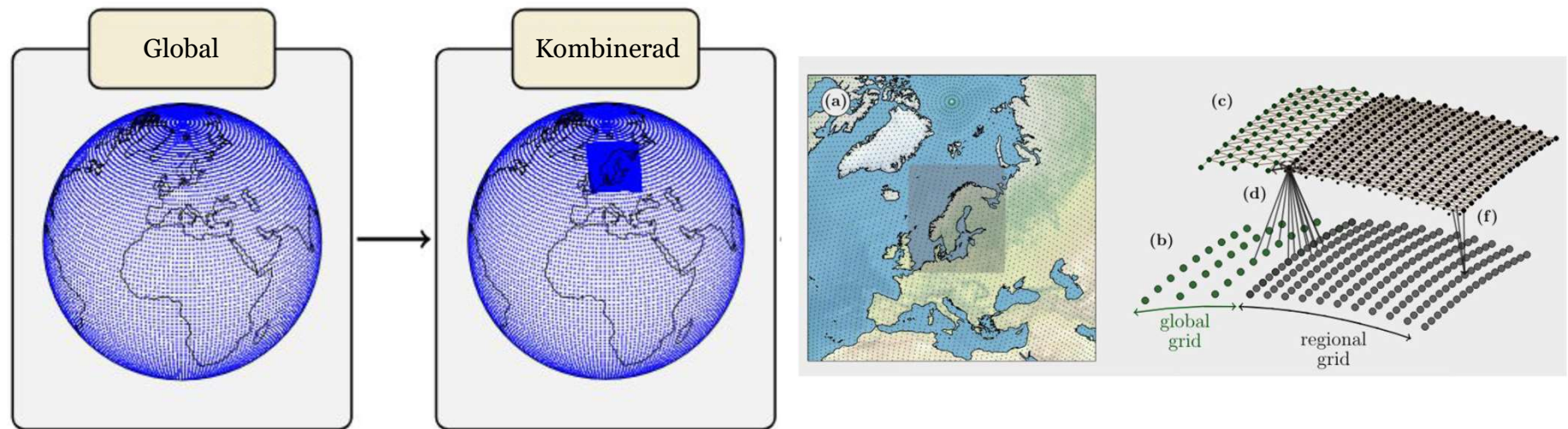
- Kan vi hantera randvillkoren på ett bättre sätt?
- Kombinera information från global och lokal modell
- Maskininlärning är flexibelt!



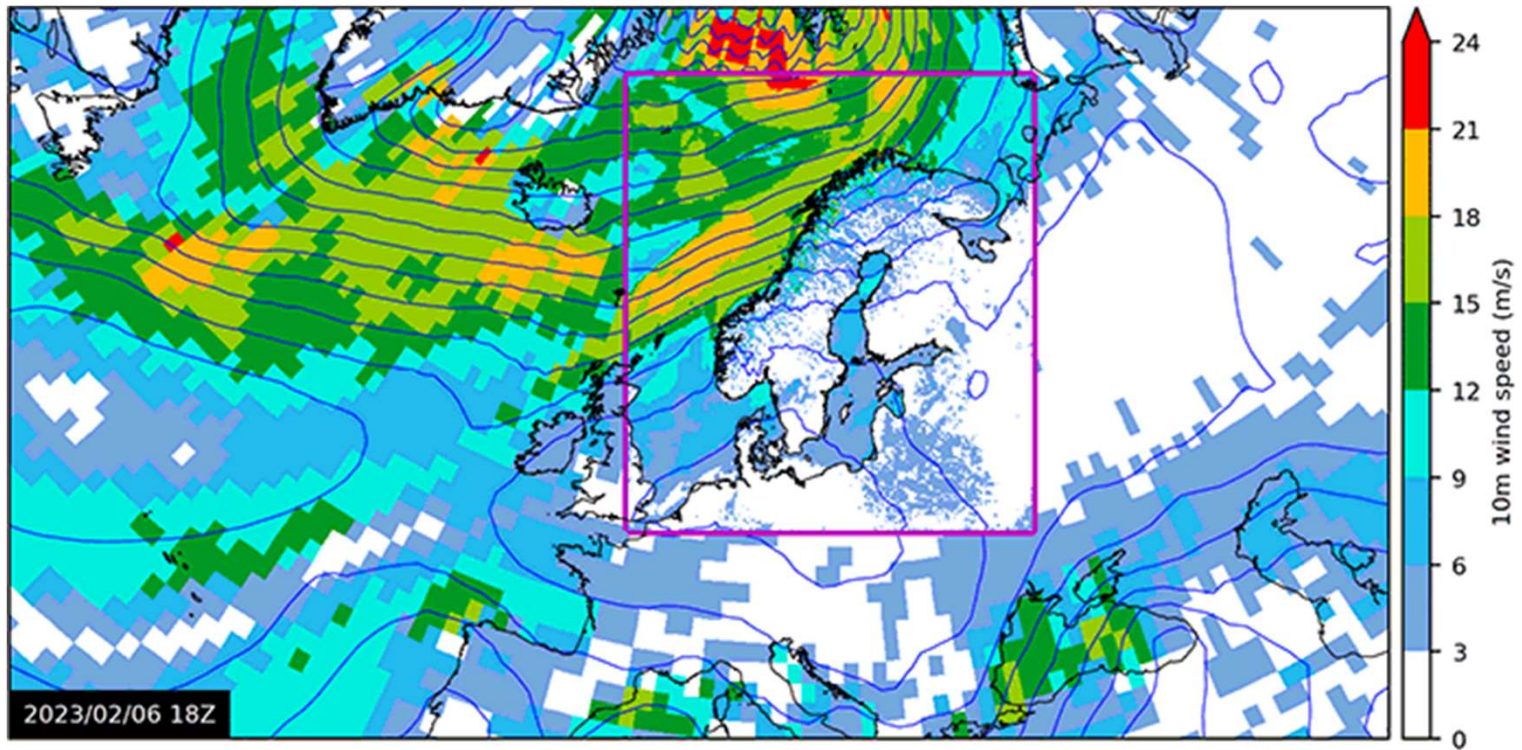
² T. Nipen, et al. (2024). Regional Data-driven modelling with a global stretched grid. *arXiv preprint*.

Global Stretched Grid

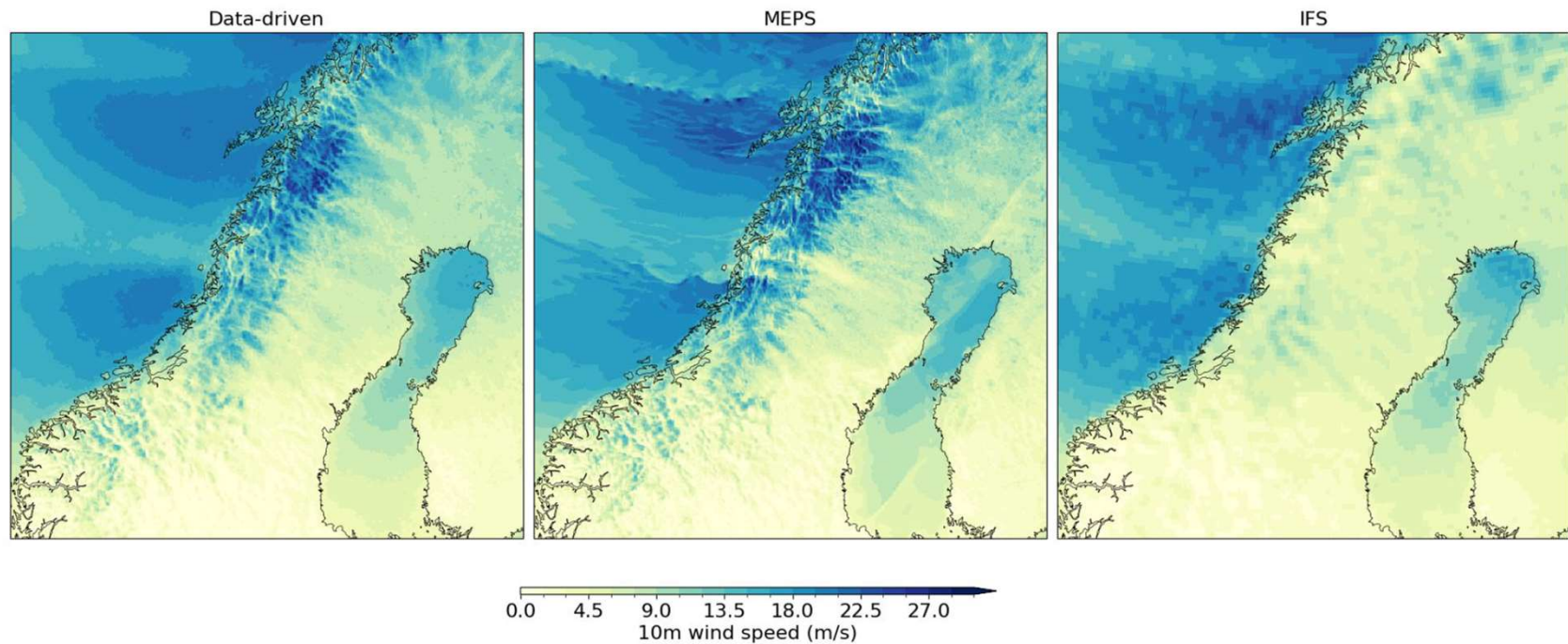
- En global modell med extra upplösning över det lokala området
- Globala randvillkor redan inbyggda i modellen



Exempelprognoser



Prognos 1 dag Vindhastighet 10m

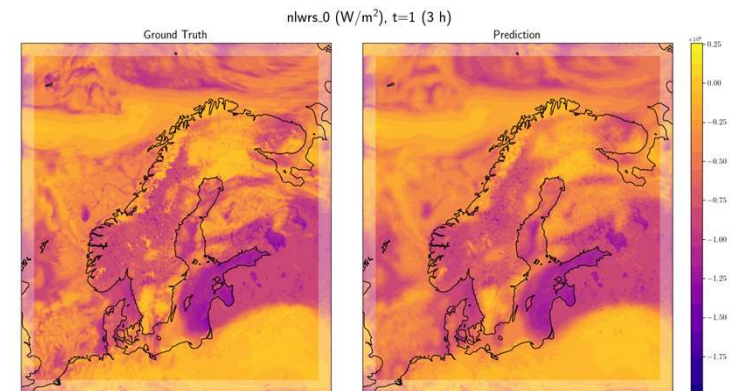


Ensembleprognoser

- Hur gör vi ensembleprognoser med maskininlärning?
- Hur kan vi utnyttja att vi har en maskininlärningsmodell?

Ensembleprognoser mha perturbering

- Prognoserna är 100x-1000x gånger snabbare att köra
- Gör som i NWP och perturbera initialvillkoren/ modellparametrar
- Deterministisk MLWP ger suddiga prognoser
-> Suddiga medlemmar
- Problemet måste behandlas med probabilistiska modeller



Generativa modeller

- Vi har generativa modeller för:
 - Text, Bilder, Ljud, Video
 - ChatGPT, Stable Diffusion, SORA etc

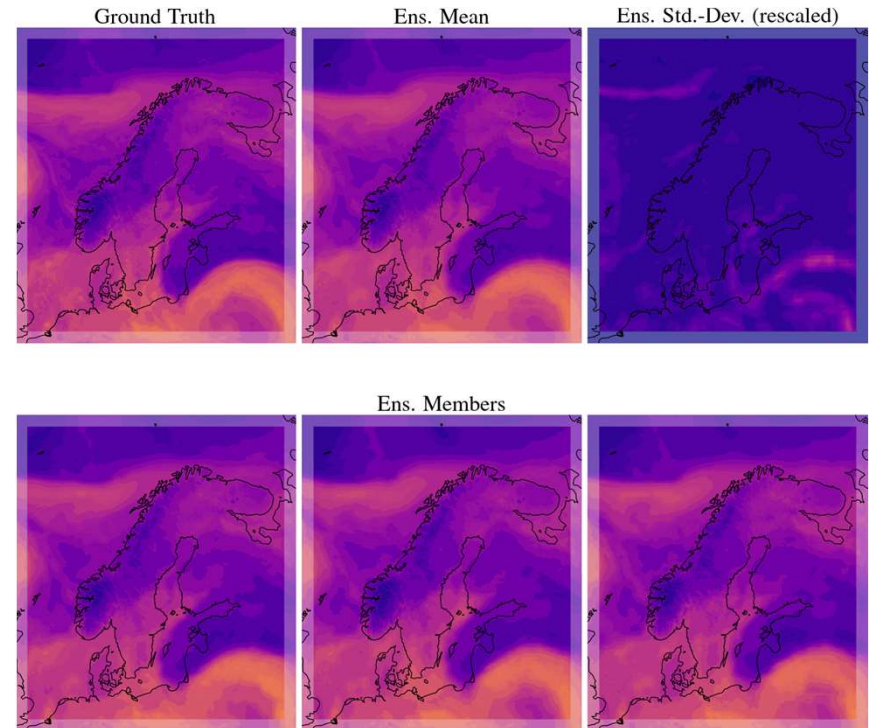
AI-genererad Video från OpenAI Sora



- Det underliggande problemet är samma som i väder
- Vi lär oss att dra från en betingad fördelning
 - $p(\text{video} | \text{textbeskrivning av videon})$
 - $p(\text{vädret imorgon} | \text{vädret idag})$

Exempel: Graph-EFM ³

- **Mål:** Kan vi emulera MEPS bättre med generativ ML?
- Enskilda medlemmar har högre upplösning
- Meningsfull osäkerhetsestimering



Vattenånga (*wvint*)

Exempel: Min forskning

Exempel på de möjligheter som finns med ML för väder

NWP

- Ensembler genom perturbering
- Stabilitetskrav på tidssteg
- Prognoser görs iterativt

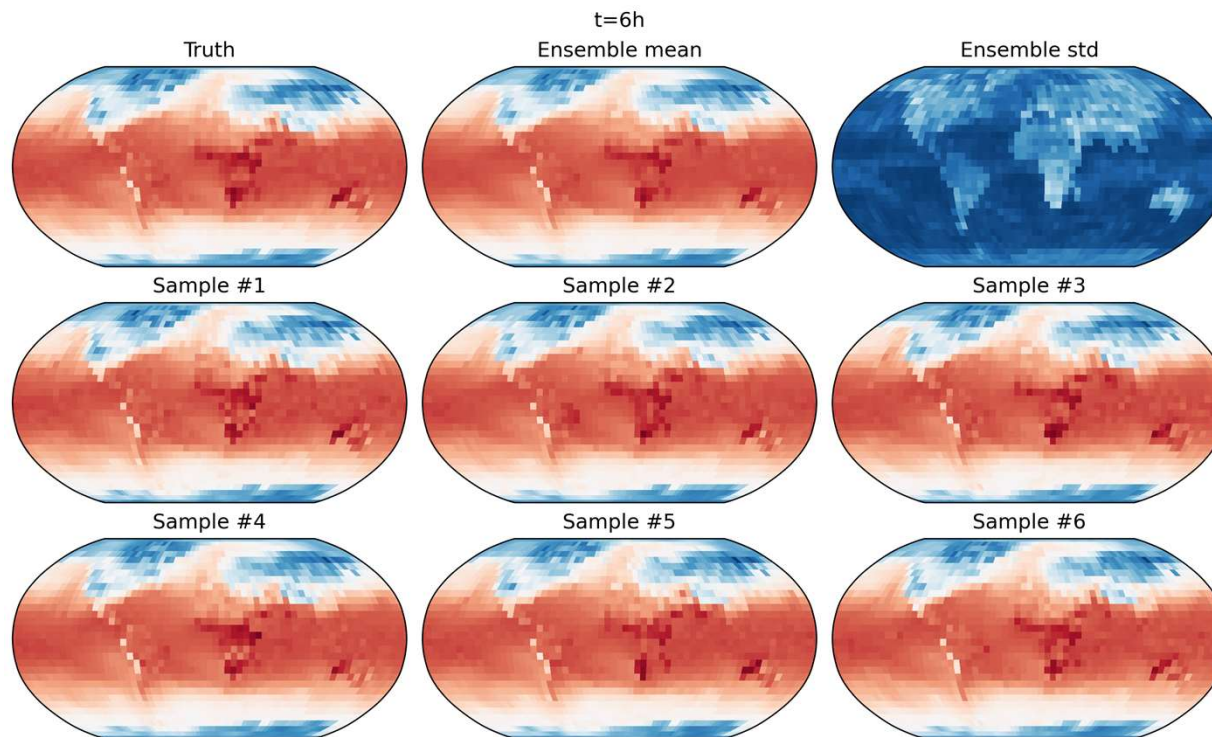
MLWP

- Osäkerhet inbyggt i modellen
- Inga explicita stabilitetskrav
- Möjligt att göra en prognos för flera dygn framåt i ett steg

Exempel: Min forskning

- 5 dygnsprognos med 1 timmars intervall á 50 ensemblemedlemmar
 - Totalt $120 \times 50 = 6000$ prognoser
 - Om en prognos tar 1 minut tar det **100h** att göra alla prognoser
- Vi kan parallellisera ensembleprognoserna effektivt på GPUer
 - Totalt **2h** för att göra alla prognoser
- Men flaskhalsen kommer fortfarande vara de iterativa stegen
 - Om vi hade kunnat parallellisera hela prognosen hade det tagit oss totalt **1 minut** att göra alla prognoser

Med ML är detta möjligt!



Prognoser av marktemperatur genererade helt utan iteration

Avslutande tankar

- Forskningen går fort och gynnas av utvecklingen i andra fält
- Möjligheterna med ML för väder är enorma
 - ~100x gånger snabbare prognoser
 - Möjligt att ha fler ensemblemedlemmar
 - Bättre täckning av extremväder
 - Flexibelt och skalbart
 - Anpassat för att köras på vilken hårdvara som helst
 - Möjligt att skräddarsy prognosområde och upplösning efter kund on demand

Tack så mycket!

martin.andrae@liu.se